

Efficient Control of Epidemics over Random Networks

Marc Lelarge
INRIA-ENS
Paris, France
Marc.Lelarge@ens.fr

ABSTRACT

Motivated by the modeling of the spread of viruses or epidemics with coordination among agents, we introduce a new model generalizing both the basic contact model and the bootstrap percolation. We analyze this percolated threshold model when the underlying network is a random graph with fixed degree distribution. Our main results unify many results in the random graphs literature. In particular, we provide a necessary and sufficient condition under which a single node can trigger a large cascade. Then we quantify the possible impact of an attacker against a degree based vaccination and an acquaintance vaccination. We define a security metric allowing to compare the different vaccinations. The acquaintance vaccination requires no knowledge of the node degrees or any other global information and is shown to be much more efficient than the uniform vaccination in all cases.

Categories and Subject Descriptors

G.2.2 [Discrete Mathematics]: Graph Theory—*Network problems*

General Terms

Theory, Performance, Design

Keywords

Epidemics, Random graphs, Vaccination

1. INTRODUCTION

Many network phenomena are well modeled as spreads of epidemics through a network. Examples include the spread of worms and email viruses, and, more generally, faults. The propagation of information and opinions can also often be modeled as the spread of an epidemic. Interacting particle systems (like the contact process) have been successfully employed to model such diverse phenomena. The distinguishing features of these stochastic systems are that they model

a large number of particles on a graph, and that each particle is assigned a state function of the states of nearby particles. Here particles may represent computers, people... Traditionally, research on interacting particle systems has only answered analysis-type questions, such as determining the steady-state behavior of the system or the probability of an epidemic outbreak. Our aim is to study these models with an engineering eye to answer design-oriented questions. With this general aim in mind, we consider the concrete problem of containment of the spread of viruses in a random graph. We will consider vaccination strategies applicable before the epidemic starts spreading as opposed to vaccination strategies based on contact tracing method (see [4] for a rigorous treatment of this case). We show that the acquaintance vaccination proposed in [7] has the advantage of requiring no global information and allows a much better protection of the network against epidemics (propagating according to a contact process or with a probability of contagion that is a threshold function) than the uniform vaccination.

A worm is a self-propagating malicious program that exploits security vulnerabilities and does not require user action to propagate. The mechanism by which a worm discovers vulnerable hosts is one of the critical factors of a worm's strategy [31]. Current generation worms have typically used simple strategies, such as random or sequential scanning of the address space, to discover vulnerable hosts. In this case, it might be reasonable to model the spread of the epidemic as a contact process [32]: a node can be infected by its infected neighbors at a rate that is proportional to their number. There is a vast literature on epidemics on complex networks (see [25] for a review) and there is now a good understanding of the impact of the topology on the spread of an epidemic [11, 8]. In this paper, we propose a new epidemic model which generalizes the basic SIR model. Most of the epidemic models consider a transmission mechanism which is independent of the local condition faced by the agents concerned. However, if there is a factor of coordination or persuasion involved, relative considerations tend to be important in understanding whether some new belief or behavior is adopted. In such cases, the probability that a node becomes infected when r out of her d neighbors are already infected can be modeled by a threshold function: the probability is $v \in (0, 1]$ if $r/d \geq \theta$ and it is zero otherwise. If $\theta = 1/2$, this would correspond to a local majority rule [27]. This notion has been proposed to address the central problem in distributed computing of overcoming failures. The idea is to eliminate the damage caused by failed vertices by maintaining replicated copies of crucial data and performing

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGMETRICS/Performance'09, June 15–19, 2009, Seattle, WA, USA.

Copyright 2009 ACM 978-1-60558-511-6/09/06 ...\$5.00.

a voting process among the participating processors whenever faults occur, adopting the values stored at the majority of the processors as the correct data. [22] proposed to apply such ideas to large peer-to-peer systems. It is then imperative to understand the possible influence of malign peer attempting to subvert the system. Such an understanding can lead to better techniques for preventing and fighting this malicious intent.

Another motivation for our work comes from social sciences. In social contexts, the diffusion of information and behavior often exhibits features that do not match well those of the contact process [29]. To illustrate our purpose, consider the basic game-theoretic diffusion models proposed by Morris [24]. Consider a graph G in which the nodes are the individuals in the population and there is an edge (i, j) if i and j can interact with each other. Each node has a choice between two possible behaviors labeled A and B . On each edge (i, j) , there is an incentive for i and j to have their behaviors match, which is modeled as the following coordination game with parameter $q \in (0, 1)$: if i and j choose A (resp. B), they each receive a payoff of q (resp. $(1 - q)$); if they choose opposite strategies, then they receive a payoff of 0. Then the total payoff of a player is the sum of the payoffs with each of her neighbors. Even at this level of simplicity, the analysis of the possible equilibria when nodes apply a best-response update gives a number of qualitative insights. Jackson and Yariv [12] analyze the influence of network structure on diffusion of behavior in network games. Watts [30] shows that there can be a cascading sequence of nodes switching to B with only one node forced to play B while all others are playing A (see also [19]). In this paper, we will extend this result and go one step further: we show how it is possible to control (i.e. provoke or prevent) efficiently such kind of cascading sequence.

The first main contribution of our work is to introduce a new model of epidemic called the **percolated threshold model**. This model is versatile enough to cover both the contact model and the threshold model. It combines neighborhood effects and random transmission. Our second main contribution is an analysis of this model when the underlying network is a random graph with fixed degree distributions. Our main results on the **spread of the epidemic** for the percolated threshold model unify many results in the random graphs literature. In particular, we derive a necessary and sufficient condition under which a single node can trigger a large cascade. Although random graphs are not considered to be highly realistic models of most real-world networks, they are often used as first approximation and are a natural first choice for a sparse interaction network in the absence of any known geometry of the problem. Our third contribution is a **rigorous analysis of attack and vaccination strategies**. We give quantitative results on the possible impact of an attacker for a vaccination targeting high-degree nodes (but requiring global knowledge of the network) and a local immunization called the acquaintance vaccination [7]. We show that this vaccination outperforms the basic uniform vaccination, especially in scale-free networks.

While there is a vast literature providing non-rigorous analysis of immunization strategies for random networks [6, 7], results in the mathematics literature are very rare. The work of Britton, Janson and Martin-Löf [5] is a notable exception but restricted to the basic contact model. Our work

provides a new proof of their results and extends them to our percolated threshold model. In the contact model, random vaccination requires vaccinating a very large fraction of a computer network or population. In particular, as it was heuristically shown in [6, 7], stopping viruses spreading upon contact requires almost 100% vaccination when the degree distribution of the underlying graph has an infinite variance for any (positive) value of the conatgion probability. On the other hand, targeted vaccination of the most highly connected nodes, while effective, requires global information about the network, rendering it impractical in many cases. Following [7], we study a vaccination of random acquaintances of randomly selected nodes. In this way, the most highly connected nodes are immunized without requiring the knowledge of the network. Our analysis validates results of [7] and shows that in the contact model, the fraction of vaccinated nodes required to stop the epidemic can be lowered from almost 100% in the random vaccination's case to 25% if the acquaintance vaccination is used. More generally, we define the efficiency of a vaccination strategy as the reduction in the possible damages caused by a single infected node. In the case of power-law graphs, we show that the efficiency gain of the acquaintance vaccination is much higher than the one obtained by random vaccination.

The paper is structured as follows. We introduce our models for the graph, the epidemic and the attack and vaccination strategies in Section 2. Our main results on the spread of the epidemic as a function of the attack strategy (Theorem 4) and the impact of the vaccinations (Theorem 5) are given in Section 3. We present applications of our theorems in Section 4. In particular we derive the cascade condition and compare degree based vaccination and acquaintance vaccination. In Section 5, we give the main technical proofs of our paper. Section 6 summarizes the paper.

2. MODELS

We first present the model for the underlying graph, then the model for the epidemic process and finally the attack and vaccination strategies.

2.1 Graphs: the configuration model

Let $n \in \mathbb{N}$ and let $(d_i)_1^n = (d_i^{(n)})_1^n$ be a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even. We define a random multigraph with given degree sequence $(d_i)_1^n$, denoted by $G^*(n, (d_i)_1^n)$ by the configuration model [3]: take a set of d_i half-edges for each vertex i and combine the half-edge into pairs by a uniformly random matching of the set of all half-edges. Conditioned on the multigraph $G^*(n, (d_i)_1^n)$ being a simple graph, we obtain a uniformly distributed random graph with the given degree sequence, which we denote by $G(n, (d_i)_1^n)$ [13].

We will let $n \rightarrow \infty$ and assume that we are given $(d_i)_1^n$ satisfying the following regularity conditions [23]:

CONDITION 1. *For each n , $(d_i)_1^n = (d_i^{(n)})_1^n$ is a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even and, for some probability distribution $(p_r)_{r=0}^\infty$ independent of n ,*

$$(i) |\{i : d_i = r\}|/n \rightarrow p_r \text{ for every } r \geq 0 \text{ as } n \rightarrow \infty;$$

$$(ii) \lambda := \sum_r r p_r \in (0, \infty);$$

$$(iii) \sum_{i=1}^n d_i/n \rightarrow \lambda \text{ as } n \rightarrow \infty;$$

The results of this work can be applied to some other random graphs models too by conditioning on the vertex degrees. For example, for the Erdős-Rényi graph $G(n, p)$ with $np \rightarrow \lambda \in (0, \infty)$, the assumptions hold with (p_r) a Poisson distribution with mean λ .

2.2 Epidemic: percolated threshold model

Symmetric threshold model: We first present the symmetric threshold model which generalizes the contagion model of [24]: a node becomes active when a certain threshold fraction of neighbors are already active. We allow the threshold fraction to be a random variable with distribution depending on the degree of the node and such that thresholds are independent among nodes. Formally, we define for each $d \in \mathbb{N}$, a sequence of i.i.d. random variables in \mathbb{N} denoted by $(K_i(d))_{i=1}^\infty$. The threshold associated to node i is $k_i = K_i(d_i)$ where d_i is the degree of node i . We will use the notation $(\mathbf{d}, \mathbf{k}) = (d_i, k_i)_{i=1}^n$.

Now the progressive dynamic of the epidemic operates as follows: some set of nodes S starts out being active; all other nodes are inactive. Time operates in discrete steps $t = 1, 2, 3, \dots$. At a given time t , any inactive node i becomes active if its fraction of active neighbors exceeds its threshold $K_i(d_i) + 1$. This in turn may cause other nodes to become active. We will suppose that $K_i(1) = 0$ for all i , so that any leaf of the network is active as soon as her neighbor becomes active.

It is easy to see that the final set of active nodes (after n time steps if the network is of size n) only depends on the initial set S (and not on the order of the activations) and can be obtained as follows: set $X_i = \mathbf{1}(i \in S)$ for all i . Then as long as there exists i such that $\sum_{j \sim i} X_j > k_i$, set $X_i = 1$, where $j \sim i$ means that i and j share an edge in G . When this algorithm finishes, the final state of node i is represented by X_i : $X_i = 1$ if node i is active and $X_i = 0$ otherwise.

Note that we allow for the possibility $d_i < K_i(d_i)$ in which case, node i is never activated unless it belongs to the set S .

We now introduce a general model of epidemic combining neighborhood effects and random transmission.

Percolated threshold model: this model depends on a parameter $\pi \in [0, 1]$ and a distribution of random thresholds $(K(d))_{d \in \mathbb{N}}$. Given any graph G and initial set S , we now proceed in two phases.

- bond percolation: randomly delete each edge with probability $1 - \pi$ independently of all other edges. We denote the resulting random graph by G_π ;
- apply the symmetric threshold epidemic with thresholds $K(d)$: set $X_i = \mathbf{1}(i \in S)$ and then as long as there is i such that $\sum_{j \sim_\pi i} X_j > K_i(d_i)$, set $X_i = 1$, where $j \sim_\pi i$ means that i and j share an edge in G_π and d_i is the degree of node i in the original graph G .

The final fraction of active nodes is simply $n^{-1} \sum_{i=1}^n X_i$.

Clearly if $\pi = 1$, this is exactly the symmetric threshold model. In general, the intuition is the following: in a case where the virus propagates instantaneously but only if some interactions occur between the nodes. If one looks at the epidemic on a finite time horizon and two neighbors did not interact during that time, the virus will not propagate on their shared edge. Hence one can think of π as the probability that they actually did interact during this period of time.

Examples of epidemics. We now present some examples to illustrate the scope of our framework.

EXAMPLE 1. *The most basic epidemic model is the contact model with probability of contagion π on each edge. It corresponds to the percolated threshold model with parameter π and where $K(d) = 0$ for any $d \in \mathbb{N}$. We refer to [18] for a connection with the dynamic SIR epidemic model.*

EXAMPLE 2. *The linear threshold model generalizes the model of [24] by allowing nodes to weigh the influence of their neighbors differently and assuming that each node's threshold is chosen randomly and possibly depends on the degree of the node. Following [16], we introduce*

- for each $d \in \mathbb{N}$, let $(W_{1i}(d), \dots, W_{di}(d))_{i=1}^\infty$ be a sequence of i.i.d. random variables such that the components of $(W_{1i}(d), \dots, W_{di}(d))$ are exchangeable;
- for each $d \in \mathbb{N}$, let $(\tilde{K}_i(d))_{i=1}^\infty$ be a sequence of i.i.d. random variables in \mathbb{R} .

For node i with degree d_i , the non-negative weight $W_{ki}(d_i)$ indicates the influence that its k 's neighbor exerts on i and the threshold $\tilde{K}_i(d_i)$ indicates the weighted fraction of i 's neighbors that must adopt the behavior before i does. Given a node i of degree d_i and a bijective mapping v from the set of neighbors of i to $\{1, 2, \dots, d_i\}$, if the set of active nodes is given by $X_j = 1$ for $j \neq i$, then i becomes active if $\sum_{j \sim i} X_j W_{v(j)i}(d_i) > \tilde{K}_i(d_i)$. Thanks to exchangeability, we have

$$\begin{aligned} \mathbb{P} \left(\sum_{j \sim i} X_j W_{v(j)i}(d_i) > \tilde{K}_i(d_i) \mid \sum_{j \sim i} X_j = k \right) \\ = \mathbb{P} \left(\sum_{\ell=1}^k W_{\ell i}(d_i) > \tilde{K}_i(d_i) \right). \end{aligned}$$

Hence this model corresponds to our general threshold model with $\mathbb{P}(K_i(d) \leq k) = \mathbb{P}(\sum_{\ell=1}^k W_{\ell 1}(d) > \tilde{K}_1(d))$.

2.3 Attack and Vaccination strategies

We first describe the vaccination strategy. We assume that a perfect vaccine is available. By this, we mean that an individual who is vaccinated is completely protected from the virus and is not able to spread the virus further. She remains inactive for the whole period of time. We assume that a part of the population is vaccinated before the epidemic starts. The epidemic progresses as defined above with the only difference that activated individuals can only activate a neighbor who has not been vaccinated. For the study of the epidemic in the vaccinated population, we may simply remove all vaccinated individuals from G (and edges connected to these individuals) but keeping the same thresholds $(k_i)_1^n$. We will consider the two following vaccination strategies:

Degree based vaccination: randomly inoculate vertex i with probability $\beta_{d_i}^P$, where d_i is the degree of the vertex.

In order to select high-degree nodes (as explained above), the whole graph (or at least the degrees of all nodes) has to be known and this is rarely the case [25]. Our next strategy aims at targeting nodes with high degree but only uses local graph-knowledge from selected nodes. The main underlying idea (first proposed in [7]) is to use the following bias: an edge points to a node with 'larger' degree in the sense

that its degree distribution is rp_r/λ , the size-biased degree distribution.

Acquaintance vaccination: sample each node with probability c' . If the node is of degree d , inoculate each of her neighbors (in the graph G) at random with probability $p'_d \leq 1$ (independently from each other). Consequently, a node with degree d will target and vaccinate each of her neighbors with probability $\beta_d^A := c'p'_d$.

We now describe the attack. In our model, the initial set of active nodes S corresponds to the attacked nodes: these nodes will remain active during the whole process of the infection. We will suppose that the attacker is able to detect vaccinated nodes and do not attack them. In other words, the initial attack (and the epidemic) takes place on the graph from which we removed the vaccinated nodes. The initial set S of active nodes depends on the attacker strategy. We will look at the following attack strategies:

Degree based attack: randomly attack vertex i with probability α_{d_i} , where d_i is the degree of the vertex. In other words, each non-vaccinated node i draws independently of each other a Bernoulli random variable σ_i with parameter α_{d_i} and is considered as active if $\sigma_i = 1$ and not initially active otherwise.

In particular, if $\alpha_d = \alpha$ for all d , then the attacker successfully contaminates a fraction α chosen uniformly in the population (before the epidemic takes place). It is intuitively clear that a better attack strategy would be to contaminate the individuals with the highest degrees. This case would correspond to a parameter $\alpha_d \geq \alpha_{d'}$ whenever $d \geq d'$, but it requires a global knowledge of the graph.

3. MAIN RESULTS

In this section, we present our main technical results. The practical insights about the interplay between attack and vaccinations follow from these results and will be described in Section 4 (which can be read while skipping this section). In Section 3.1, we show how to modify the parameter of the degree distribution of the graph when vaccinated nodes are removed. In Section 3.2, we compute the size of the epidemic on a general graph for a degree based attack. By applying this last result to the modified graph (where vaccinated nodes are removed), we can measure the interplay between the attack and vaccination strategies. Finally in Section 3.3, we provide more physical meanings for our formula thanks to the local mean field approximation (shown to be exact in our case in [17]).

We first need to introduce some notations. For $d, \ell \in \mathbb{N}$, we denote $t_{d\ell} = \mathbb{P}(K(d) = \ell)$ the probability distribution of the threshold for a node of degree d . For integers $\ell \geq 0$ and $0 \leq r \leq \ell$ let $b_{\ell r}$ denote the binomial probabilities $b_{\ell r}(p) := \mathbb{P}(\text{Bi}(\ell, p) = r) = \binom{\ell}{r} p^r (1-p)^{\ell-r}$.

All unspecified limits and other asymptotics statement are for $n \rightarrow \infty$. For example, w.h.p. (with high probability) means with probability tending to 1 as $n \rightarrow \infty$ and \rightarrow^p means convergence in probability as $n \rightarrow \infty$. Similarly, we use o_p and O_p in a standard way. For example, if X is a parameter of the random graph, $X = o_p(n)$ means that $\mathbb{P}(X > \epsilon n) \rightarrow 0$ as $n \rightarrow \infty$ for every $\epsilon > 0$; equivalently $X/n \rightarrow^p 0$.

3.1 Vaccination

The vectors $\alpha = (\alpha_d)_{d=1}^\infty$, $\beta^P = (\beta_d^P)_{d=1}^\infty$ and $\beta^A = (\beta_d^A)_{d=1}^\infty$ are the parameters for the attack strategy, the de-

gree based vaccination strategy and the acquaintance vaccination strategy respectively (introduced in Section 2.3).

PROPOSITION 2. Consider the random graph $G^*(n, (d_i, k_i)_1^n)$ for a sequence $(d_i)_1^n$ satisfying Condition 1.

Let $G^*(n^P, (d_i^P, k_i^P)_1^{n^P})$ (resp. $G^*(n^A, (d_i^A, k_i^A)_1^{n^A})$) be the random graph obtained when removing all vaccinated nodes from $G^*(n, (d_i, k_i)_1^n)$ for the degree based (res. acquaintance) vaccination with parameter β^P (resp. β^A).

For the degree based vaccination, we have:

- $n^P = \zeta^P n + o_p(n)$ with $\zeta^P = \sum_j (1 - \beta_j^P) p_j$;
- let $x = \frac{1}{\lambda} \sum_{s \geq 1} s(1 - \beta_s^P) p_s$, then we have

$$\begin{aligned} & \frac{|\{i : d_i^P = j; k_i^P = \ell\}|}{n^P} \\ & \rightarrow^p p_{j\ell}^P = \frac{1}{\zeta^P} \sum_{s \geq j} (1 - \beta_s^P) p_s t_{s\ell} b_{sj}(x). \quad (1) \end{aligned}$$

For the acquaintance vaccination, we have:

- $n^A = \zeta^A n + o_p(n)$ with $\zeta^A = \sum_\ell p_\ell y^\ell$, where $y = \frac{1}{\lambda} \sum_{s \geq 1} s(1 - \beta_s^A) p_s$;
- let $\hat{p} = \frac{\sum_\ell \ell p_\ell y^\ell (1 - \beta_\ell^A)}{\lambda y^2}$, then we have

$$\begin{aligned} & \frac{|\{i : d_i^A = j; k_i^A = \ell\}|}{n^A} \\ & \rightarrow^p p_{j\ell}^A = \frac{1}{\zeta^A} \sum_{s \geq j} y^s p_s t_{s\ell} b_{sj}((1 - \beta_s^A) \hat{p}). \quad (2) \end{aligned}$$

If in addition, $\sum_i d_i^2 = O(n)$, then all the conclusions hold also for the (simple) random graphs $G(n, (d_i, k_i)_1^n)$, $G(n^P, (d_i^P, k_i^P)_1^{n^P})$ and $G(n^A, (d_i^A, k_i^A)_1^{n^A})$.

Proposition 2 is a direct consequence of Lemmas 11 and 12 proved in Sections 5.3 and 5.4.

COROLLARY 3. For the degree based vaccination, the asymptotic fraction of vaccinated nodes is $1 - \zeta^P$ and the asymptotic degree distribution in the graph where vaccinated nodes have been deleted is

$$p_j^P = \frac{1}{\zeta^P} \sum_{s \geq j} (1 - \beta_s^P) p_s b_{sj}(x),$$

where ζ^P and x are defined in Proposition 2.

For the acquaintance vaccination, the asymptotic fraction of vaccinated nodes is $1 - \zeta^A$ and the asymptotic degree distribution in the graph where vaccinated nodes have been deleted is

$$p_j^A = \frac{1}{\zeta^A} \sum_{s \geq j} y^s p_s b_{sj}((1 - \beta_s^A) \hat{p}),$$

where ζ^A, y and \hat{p} are defined in Proposition 2.

3.2 Epidemic spread

Given a distribution $\mathbf{p} = (p_{s\ell} = p_s t_{s\ell})$, we define the functions:

$$h(z; \alpha, \mathbf{p}, \pi) := \sum_{s, r \geq s - \ell} r(1 - \alpha_s) p_s t_{sr} (1 - \pi + \pi z),$$

$$g(z; \alpha, \mathbf{p}, \pi) := h(z; \alpha, \mathbf{p}, \pi) - \lambda(\mathbf{p}) z (1 - \pi + \pi z),$$

$$h_1(z; \alpha, \mathbf{p}, \pi) := \sum_{s, r \geq s - \ell} (1 - \alpha_s) p_s t_{sr} (1 - \pi + \pi z),$$

where $\lambda(\mathbf{p}) = \sum_s s \sum_\ell p_{s\ell}$. For a parameter of the attack α , we define

$$\hat{z}(\mathbf{p}, \alpha, \pi) := \max \{z \in [0, 1] : g(z; \alpha, \mathbf{p}, \pi) = 0\}.$$

The next result allows to compute for any parameter α of the attack, the final proportion of infected nodes: $\Phi(\alpha) = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n X_i}{n}$ (defined in Section 2.2).

THEOREM 4. *Consider the random graph $G^*(n, (d_i, k_i)_1^n)$ for a sequence $(d_i)_1^n$ satisfying Condition 1 and let $\mathbf{p} = (p_{st_{s\ell}})$. For the percolated threshold epidemic on the graph $G^*(n, \mathbf{d}, \mathbf{k})$, we have:*

- (i) *If $\hat{z}(\mathbf{p}, \alpha, \pi) = 0$, then w.h.p. $\Phi(\alpha) = 1 - h_1(0, \alpha, \mathbf{p}, \pi)$.*
- (ii) *If $\hat{z}(\mathbf{p}, \alpha, \pi) \in (0, 1]$, and further \hat{z} is not a local maximum point of $g(z; \alpha, \mathbf{p}, \pi)$, then w.h.p. $\Phi(\alpha) = 1 - h_1(\hat{z}, \alpha, \mathbf{p}, \pi)$.*

If in addition, $\sum_i d_i^2 = O(n)$, then all the conclusions hold also for the (simple) random graph $G(n, \mathbf{d}, \mathbf{k})$.

A proof of this theorem is given in [17]. Comparison with existing results in the literature is given in Section 4.

We now state our main results concerning the interaction between vaccination and attack strategies.

THEOREM 5. *For the degree based vaccination, the conclusions of Theorem 4 hold for the random graph $G^*(n^P, (d_i^P, k_i^P)_1^n)$ with the parameter $\mathbf{p}^P = (p_{j\ell}^P)$ defined in (1).*

For the acquaintance vaccination, the conclusions of Theorem 4 hold for the graph $G^(n^A, (d_i^A, k_i^A)_1^n)$ with the parameter $\mathbf{p}^A = (p_{j\ell}^A)$ defined in (2).*

In words, Theorem 5 shows that for the percolated threshold epidemic, the graphs obtained by deletion of the vaccinated nodes, behave like random graphs obtained by the configuration model with the appropriate asymptotic degree-threshold sequence. The proof of Theorem 5 follows directly from Theorem 4 and Lemmas 11 and 12 (proved in Section 5).

3.3 Local Mean Field approximation

In this subsection, we describe an approximation to the local structure of the graph and present a heuristic argument which leads quickly to a prediction for the asymptotic probability of being active which is verified rigorously in [17]. This branching process approximation is standard in the random graphs literature [9] and we called it Local Mean Field (LMF) in [21, 20] for contact epidemics. In this case, we were able to turn the LMF approximation into a rigorous argument. For our percolated threshold model, this seems unlikely to be straightforward and our proof will not use the LMF approximation. However, for certain random graphs, this sort of approach was made to work as in [28].

The LMF model is characterized by its connectivity distribution p_r . We now construct a tree-indexed process. Let T be a Galton-Watson branching process with a root which has offspring distribution p_r and all other nodes have offspring distribution p_r^* given by $p_{r-1}^* = \frac{r p_r}{\lambda}$ for all $r \geq 1$. Recall that p_{r-1}^* corresponds to the probability that an edge points to a node with degree r [9]. The tree T describes the local structure of the graph G (as n tends to infinity): the exploration of the successive neighbourhoods of a given

vertex is approximated by the branching process T as long as the exploration is local (typically restricted to a finite neighbourhoods independent of n).

We denote by \circ the root of the tree and for a node i , we denote by $\text{gen}(i) \in \mathbb{N}$ the generation of i , i.e. the length of the minimal path from \circ to i . Also we denote $i \rightarrow j$ if i belongs to the children of j , i.e. $\text{gen}(i) = \text{gen}(j) + 1$ and j is on the minimal path from \circ to i . For an edge (i, j) with $i \rightarrow j$, we denote by $T_{i \rightarrow j}$ the sub-tree of T with root i obtained by the deletion of edge (i, j) from T .

We now consider the percolated threshold model on the tree T , where the initial set of active nodes is given by a vector $\sigma = (\sigma_i)$ as described in Section 2.3. Thanks to the tree structure, it is possible to compute the probability of being active recursively as follows: for any node $i \neq \circ$, let $Y_i = 1$ if node i is active on the sub-graph $T_{i \rightarrow j} \cup \{(i, j)\}$ with initial set of active nodes given by the restriction of σ to individuals in $T_{i \rightarrow j}$ and with individual j held fix in the inactive state. Then for any $i \neq \circ$, we get

$$Y_i = 1 - (1 - \sigma_i) \mathbb{1} \left(\sum_{\ell \rightarrow i} B_{\ell i} Y_\ell \leq K(d_i) \right), \quad (3)$$

where the $B_{\ell i}$'s are independent Bernoulli random variables with parameter π . Then the state of the root is given by

$$X_\circ = 1 - (1 - \sigma_\circ) \mathbb{1} \left(\sum_{i \rightarrow \circ} B_{i \circ} Y_i \leq K(d_\circ) \right). \quad (4)$$

In order to compute the distribution of X_\circ , we first solve the Recursive Distributional Equation (RDE) associated to the Y_i 's: thanks to the tree structure, the random variables Y_ℓ in (3) are i.i.d. and have the same distribution as Y_i . Hence their distribution solve the RDE given by

$$Y \stackrel{d}{=} 1 - (1 - \sigma(D^* + 1)) \mathbb{1} \left(\sum_{i=1}^{D^*} B_i Y_i \leq K(D^* + 1) \right), \quad (5)$$

where for a given d , the random variable $K(d)$ is defined in Section 2.2, $\sigma(d)$ is Bernoulli with parameter α_d , B_i 's are independent Bernoulli with parameter π , D^* has distribution p_r^* , Y and the Y_i are i.i.d. copies (with unknown distribution).

LEMMA 6. *Let $x = \mathbb{P}[Y = 1]$, where the distribution of Y solves the RDE (5), then we have $\lambda(1-x)(1-x\pi) = h(1-x; \alpha, \mathbf{p}, \pi)$. Moreover X_\circ in (4) is a Bernoulli random variable with parameter $1 - h_1(1-x; \alpha, \mathbf{p}, \pi)$.*

PROOF. Taking expectation in (5), we get

$$\begin{aligned} 1 - x &= \sum_{s \geq 0} p_s^* (1 - \alpha_{s+1}) \mathbb{P} \left(\sum_{i=1}^s B_i Y_i \leq K(s+1) \right) \\ &= \left\{ (1 - \alpha_1) \frac{p_1}{\lambda} \right. \\ &\quad \left. + \sum_{s \geq 1} \frac{(s+1)p_{s+1}}{\lambda} (1 - \alpha_{s+1}) \sum_{\ell} t_{s+1\ell} \sum_{j \leq \ell} b_{sj}(x\pi) \right\} \\ &= \frac{1}{\lambda} \left\{ (1 - \alpha_1) p_1 \right. \\ &\quad \left. + \sum_{s \geq 1; j \geq s-\ell} (1 - \alpha_{s+1}) p_{s+1} t_{s+1\ell} (s+1) b_{sj} (1 - x\pi) \right\}. \end{aligned}$$

Note that $(s+1)b_{sj}(1-x\pi) = \frac{j+1}{1-x\pi}b_{s+1,j+1}(1-x\pi)$ for $s \geq 1$, so that

$$\begin{aligned} 1-x &= \frac{1}{\lambda} \left\{ (1-\alpha_1)p_1 \right. \\ &+ \left. \sum_{s \geq 1; j \geq s-\ell} \frac{(j+1)}{1-x\pi} (1-\alpha_{s+1})p_{s+1\ell}b_{s+1,j+1}(1-x\pi) \right\} \\ &= \frac{1}{\lambda(1-x\pi)} \sum_{s \geq 1; j \geq s-\ell} j(1-\alpha_s)p_{s\ell}b_{sj}(1-x\pi), \end{aligned}$$

where in the last equality we used the fact that for $s=1$, we have $t_{1\ell} = \mathbf{1}(\ell=0)$. Hence, we get

$$(1-x\pi)(1-x)\lambda = \sum_{s \geq 1; j \geq s-\ell} j(1-\alpha_s)p_{s\ell}b_{sj}(1-x\pi).$$

This establishes the first part of the lemma. Taking expectation in (4) gives: $\mathbb{E}[X_\circ] = 1 - \sum_{s; j \geq s-\ell} (1-\alpha_s)p_{s\ell}t_{s\ell}b_{sj}(1-x\pi)$, and the second part of the lemma follows. \square

By the change of variable $z = 1-x$, we see that Lemma 6 is consistent with Theorem 4. Clearly the crucial point in recursion (3) is the fact that the Y_i can be computed ‘‘bottom-up’’. The Y_i ’s in (4) encode the information that i is activated by a node in the subtree of T ‘‘below’’ i (and not by the root). If one considers a node in the original graph and runs a directed contagion model on a local neighborhood of this node where only ‘directed’ contagion toward this node are allowed, then the state of the graph seen from this node is well approximated by the Y_i ’s.

4. APPLICATIONS

We now show how our Theorems 4 and 5 give new results and insights for vaccination strategies. We first derive the cascade condition (rigorously establishing a result of [30]): if the parameters of the original graph and epidemics satisfy this condition, a single active node can trigger a global cascade. We then analyze the impact of the vaccinations and compare the efficiency of the various strategies. In the last two subsections, we show that our Theorem 4 allows to recover standard results about random graphs.

4.1 Cascade condition

We now look at the possible damages caused by a small attack, made on a negligible fraction of the population. In our model it corresponds to the case where $\alpha \rightarrow 0$. Let D be a random variable with distribution $\mathbb{P}(D=r) = p_r$. We call the following condition the cascade condition:

$$\pi \mathbb{E}[D(D-1)\mathbf{1}(K(D)=0)] > \mathbb{E}[D]. \quad (6)$$

PROPOSITION 7. *If the cascade condition (6) is satisfied, then a single initial active node can activate a strictly positive fraction (as $n \rightarrow \infty$) of the population for the percolated threshold model. If the condition is not satisfied, then for any finite set S of initially active nodes, the final fraction of active nodes is $o_p(n)$.*

Before proving this proposition, we make some remarks:

- note that if K does not depend on d , then the condition becomes

$$\pi \mathbb{P}(K=0) > \frac{\mathbb{E}[D]}{\mathbb{E}[D(D-1)]}.$$

In particular, if $K \equiv 0$, then we find the well-known epidemic threshold condition for the contact model (see (9) below).

- in the case, where $K(d) = qd$ and $\pi = 1$, [19]. Watts obtained the same condition by an heuristic argument validated through simulations (see the cascade condition Eq. 5 in [30]).

PROOF. We consider the case $\alpha_r = \alpha$ and introduce the following threshold: $K_i(d) = (d+1)\mathbf{1}(K_i(d_i) \geq 1)$. In words, a node i becomes active if one of his neighbor is active and $K_i(d_i) = 0$. Clearly the nodes that become active in this model need to have only one active neighbor in the original contagion model with parameter $K_j(d_j)$. For any node i , let C_i be the final set of active nodes when starting with only i as active node. Clearly, if $j \in C_i$, then we have $C_i \setminus \{i\} \subset C_j$. Now if we prove that $\Phi(0+) = \lim_{\alpha \rightarrow 0} \Phi(\alpha) > 0$, then it is easy to show that there is a node i such that $|C_i| \sim \Phi(0+)n$ (see [17] for details). Then any point in C_i will activate at least the set $C_i \setminus \{i\}$ in the original contagion model. We now prove that for under condition (6), we have $\Phi(0+) > 0$. This will implies the first statement of the proposition. We have

$$\begin{aligned} h(z; \alpha) &= h(z; \alpha, \mathbf{p}, \pi) \\ &= (1-\alpha) \left(\sum_s s p_s t_{s0} (1-\pi + \pi z)^s \right. \\ &\quad \left. + \sum_s (1-t_{s0}) s p_s (1-\pi + \pi z) \right). \end{aligned}$$

Let $\phi(z) = \lambda z(1-\pi + \pi z) - h(z; 0)$. The condition $\Phi(0+) > 0$ is equivalent to for $\epsilon > 0$ small enough $\phi(1-\epsilon) > 0$. A simple computation shows that

$$\phi(1-\epsilon) = \epsilon \left(-\lambda + \pi \sum_s s(s-1)t_{s0}p_s \right) + o(\epsilon),$$

which is the condition of the proposition.

We need now to prove that if (6) is not satisfied, then a single active node cannot activate a positive fraction of the population. We now consider the original threshold model with $\alpha_r = \alpha$. First note that as $\epsilon \rightarrow 0$, we have for $\ell \geq 1$,

$$\sum_{r \geq s-\ell} r b_{sr}(1-\epsilon\pi) \sim s(1-\epsilon\pi).$$

Hence, with the same notation as above, we have as $\epsilon \rightarrow 0$,

$$\begin{aligned} h(1-\epsilon; \alpha) &\sim (1-\alpha) \left(\sum_s s p_{s0} (1-s\epsilon\pi) \right. \\ &\quad \left. + \sum_{s, \ell \geq 1} s(1-\pi\epsilon)p_{s\ell} \right) \\ &= (1-\alpha) \left(\lambda - \pi\epsilon(\lambda + \sum_s s(s-1)p_s t_{s0}) \right) \end{aligned}$$

so that we get

$$\phi(1-\epsilon) \sim \epsilon \left(-\lambda + \pi \sum_s s(s-1)t_{s0}p_s \right)$$

Hence we see that if (6) is not satisfied, then we have $\lim_{\alpha \rightarrow 0} \Phi(\alpha) = 0$ which implies the claim. \square

4.2 Impact of vaccination

We now compare the degree based vaccination and the acquaintance vaccination. To do this, we compute the impact of each vaccination on the cascade condition (6).

PROPOSITION 8. *For the degree based vaccination, the cascade condition becomes:*

$$\pi \sum_s s(s-1)(1-\beta_s^P)t_{s0} \frac{p_s}{\lambda} > 1, \quad (7)$$

where $t_{s0} = \mathbb{P}(K(s) = 0)$. For the acquaintance vaccination, the cascade condition becomes:

$$\pi \sum_s s(s-1)(1-\beta_s^A)^2 y^{s-2} t_{s0} \frac{p_s}{\lambda} > 1. \quad (8)$$

In the SIR case, where $K(d) = 0$, i.e. $t_{s0} = 1$ and for a choice $\beta_s^A = 1 - e^{-c/s}$ (explained below), our condition (8) corresponds to Equation (9) of [7] or Equation (3.12) of [5].

PROOF. This proposition follows from simple computations based on Lemmas 11 and 12. \square

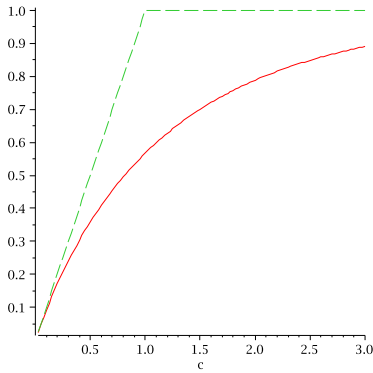


Figure 1: Fraction of vaccinated population in the acquaintance vaccination strategy $\text{vac}(c)$ (red-solid line) compared to the curve $c \mapsto \min(c, 1)$ (dashed-green line) for $\gamma = 3.2$ and $N = 5$.

From now on, we consider the following acquaintance vaccination: we suppose that each individual is sampled $\text{Po}(c)$ times, where $\text{Po}(c)$ is a Poisson distribution with parameter $c \in [0, \infty)$. Each time she reports a randomly chosen friend who is vaccinated. For simplicity, we assume that each individual does this with replacement. Consequently, this vaccination scheme corresponds to the acquaintance vaccination with $\beta_d^A = 1 - e^{-c/d}$ (studied in [7] and [5]).

All figures are done for a power-law graph where the degree distribution is given by $p_r = r^{-\gamma}/C(\gamma)$ for $r \geq N$. $N \geq 1$ and $\gamma > 2$ are two parameters and $C(\gamma) = \sum_{k \geq N} k^{-\gamma}$ is the normalizing constant.

In the acquaintance vaccination, it may happen that some individuals are chosen more than once for immunisation (being selected as neighbors of more than one individual). As a result, the fraction actually immunised denoted by $\text{vac}(c)$ is smaller than the fraction of sampled individuals. In our model each individual is sampled c times on average. Figure 1 shows the fraction of population actually vaccinated as a

function of $c \in [0, \infty)$. We see that in order to vaccinate 90% of the population, i.e. $\text{vac}(c) = 0.9$, each individual will have been randomly selected $c = 3$ times on average.

In the two following sections, we compare the impact of the uniform vaccination, i.e. the degree based vaccination with $\beta_d^P = \beta$ for all d and the acquaintance vaccination.

4.3 Vaccination for the contact model

Here we consider an epidemic modeled by a contact process, i.e. $K \equiv 0$. In particular, the cascade condition boils down to the epidemic threshold condition: if $\pi > \pi_c = \frac{\mathbb{E}[D]}{\mathbb{E}[D(D-1)]}$ then the rate of propagation is sufficiently large for the virus to become epidemic. Note in particular that if $\mathbb{E}[D^2] = \infty$, then the epidemic threshold is zero [26, 2]: viruses with very small propagation rate become epidemic.

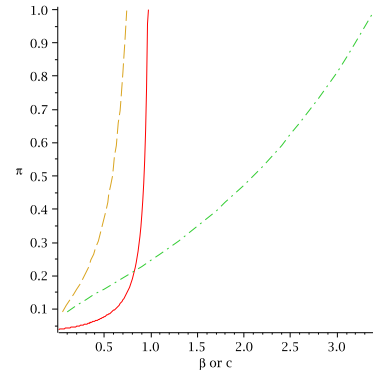


Figure 2: Epidemic threshold as a function of β for the random vaccination (red-solid), as a function of c (resp. $\text{vac}(c)$) for the acquaintance vaccination (green-dashed-dotted) (resp. (yellow-dashed)) for $\gamma = 3.2$ and $N = 5$.

From Proposition 8, we see that the uniform vaccination increases the epidemic threshold from π_c to $\pi^P(\beta) = \frac{\pi_c}{1-\beta}$ where $\beta < 1$ is the fraction of vaccinated nodes. In this case, the acquaintance vaccination is much more efficient. We rewrite (8) as follows: $\pi > \pi^A(c)$, where c is the average number of times an individual is randomly selected. Figure 2 shows the curves of the epidemic threshold for the random vaccination $\pi^P(\beta)$, for the acquaintance vaccination $\pi^A(c)$ and for the acquaintance vaccination as a function of the fraction of vaccinated population $\pi^A(\text{vac}^{-1}(\beta))$.

Note in particular that for any value of $c < 0.8$, the acquaintance vaccination strategy performs better than the uniform vaccination with parameter $\beta = c$. Even if the fraction of vaccinated nodes in the acquaintance vaccination strategy $\text{vac}(c)$ is lower than c , the efficiency of the acquaintance vaccination is much higher than the uniform vaccination because high-degree nodes have been targeted (without any global information). Here is another example of the efficiency of the acquaintance vaccination: when the acquaintance vaccination has reached 25% of the population, i.e. $\text{vac}(c) = 0.25$, the epidemic threshold is approximately 0.2 and would require almost 100% of the population being vaccinated with the uniform vaccination.

In the extreme case where $\mathbb{E}[D^2] = \infty$, the uniform vaccination is useless: it requires to vaccinate the whole population to protect it from the weakest virus (see [6]). The

acquaintance vaccination was proposed in [7] and shown (heuristically) to be able to stop a contact epidemic on a power-law graph even when the second moment of the distribution do not exist. We refer to [7] for more details. Our work (see also [5]) gives a rigorous treatment of these results.

4.4 Vaccination for the threshold model

We now consider an epidemic with threshold given by $K(D) = qD$, where q is a parameter. We examined the impact of π in the last section and we assume now that $\pi = 1$. Then q can be interpreted as a measure of the performance of the virus: the lower q is, the more contagious the virus is.

We now define the contagion threshold:

$$q_c := \sup \left\{ q : \sum_{1 \leq s < 1/q} s(s-1)p_s > \lambda \right\}.$$

The cascade condition (6) becomes (recall that $\pi = 1$): if $q < q_c$, then a global cascade is possible. We now look at the impact of the vaccinations on the contagion threshold. Namely we define

$$q^P(\beta) := \sup \left\{ q : (1-\beta) \sum_{1 \leq s < 1/q} s(s-1)p_s > \lambda \right\},$$

$$q^A(c) := \sup \left\{ q : \sum_{1 \leq s < 1/q} s(s-1)e^{-2c/s}y(c)^{s-2}p_s > \lambda \right\}.$$

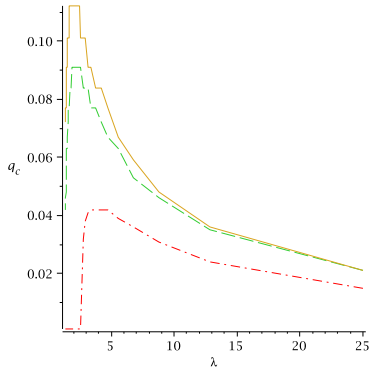


Figure 3: Curves of the contagion threshold without vaccination (q_c solid), with uniform vaccination (q_c^P dashed), with acquaintance vaccination (q_c^A dashed-dotted) as a function of λ the mean degree.

For a fixed parameter c , Figure 3 gives the curves of q_c , $q^P(1-\zeta^A)$ and $q^A(c)$ for various values of the mean degree. Hence we compare the two vaccination strategies when the asymptotic number of vaccinated nodes is the same, namely $1-\zeta^A$. We see that $q^A(c) < q^P(1-\zeta^A) < q_c$. In particular, for any $q^A(c) < q < q^P(1-\zeta^A)$, a single node can trigger a large cascade if only uniform vaccination has been done whereas he would not be able if acquaintance vaccination had been done (with the same rate of vaccination).

Clearly previous figure shows that acquaintance vaccination is more efficient than the uniform one. Figure 4 quantifies this difference and compares the effective benefit of both vaccinations. Note that the values q_c , q^P and q^A could be

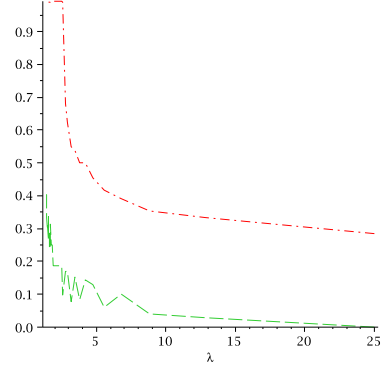


Figure 4: Efficiency gain (in %) of the uniform vaccination (dashed) and of the acquaintance vaccination (dashed-dotted) as a function of λ the mean degree.

used as a security metric (something like the inverse of the 'cost to break': the higher q_c the easier it is to attack the system). In particular q_c would be the reference and then one can compare different security solutions by computing by how much they increase the 'cost to break'. The curves on Figure 4 are given by $1-q^P/q_c$ and $1-q^A/q_c$ and hence are intrinsic measure of the efficiency of each vaccination.

4.5 Percolation in random graphs

We consider here the case where $\pi \in [0, 1]$, $K(d) = 0$ for all d and $\alpha_s = \alpha$ for all s . This model is an exploration (from the initial condition) of the connected components of the random graph after a standard bond percolation. A simple computation shows that in this case, we have with the notation of Section 3:

$$h(z; \alpha, \mathbf{p}, \pi) = (1-\alpha)(1-\pi+\pi z)g'_D(1-\pi+\pi z)$$

$$h_1(z; \alpha, \mathbf{p}, \pi) = (1-\alpha)g_D(1-\pi+\pi z),$$

where $g_D(p) = \mathbb{E}[p^D]$ is the generating function of the asymptotic degree distribution.

PROPOSITION 9. *For $\alpha \in [0, 1]$, the final fraction of active nodes is given by $\Phi(\alpha) = 1 - h_1(\hat{z}; \alpha, \mathbf{p}, \pi)$, where \hat{z} is the largest solution in $[0, 1]$ of $\lambda z = (1-\alpha)g'_D(1-\pi+\pi z)$.*

PROOF. It follows from Theorem 4 by noting that $g(z; \alpha, \mathbf{p}, \pi) = (1-\pi+\pi z)((1-\alpha)g'_D(1-\pi+\pi z) - \lambda z)$. \square

Consider now the limiting case $\alpha \rightarrow 0$. If there is a giant component in the percolated random graph, then for any $\alpha > 0$ at least one node from the initial set S will fall in this giant component and the final set of active nodes will exactly corresponds to this giant component. In this case we have:

$$g(z; 0, \mathbf{p}, \pi) = (1-\pi+\pi z)f(z), \text{ with,}$$

$$f(z) = g'_D(1-\pi+\pi z) - \lambda z.$$

In particular, we have $f(0) \geq 0$ and $f(1) = 0$ and $f'(1) = \pi\mathbb{E}[D(D-1)] - \mathbb{E}[D]$ and then, it is easy to see that if

$$\pi > \pi_c := \frac{\mathbb{E}[D]}{\mathbb{E}[D(D-1)]}, \quad (9)$$

then there is a unique solution in $\hat{z} \in (0, 1)$ such that $g(\hat{z}; 0, \mathbf{p}, \pi) = 0$ and then we have $\lim_{\alpha \rightarrow 0} \Phi(\alpha) > 0$. This corresponds to

existing results in the literature [5, 10] (see in particular Theorem 3.9 in [14]) which extend the standard result of Molloy and Reed [23] for the giant component of random graphs.

4.6 Bootstrap percolation in random regular graphs

The bootstrap percolation in a random regular graph corresponds exactly to our epidemic with parameters $p_r = 1$ for $r \geq 2$, $\pi = 1$ and $K(r) = k$, where $1 \leq k \leq r - 1$ is a fixed parameter. Since all nodes have the same degree r , we write $\alpha_r = \alpha$. In this case we have

$$\begin{aligned} h(z; \alpha, \mathbf{p}(r), 1) &= (1 - \alpha) \sum_{\ell \geq r-k} \ell b_{r,\ell}(z) \\ &= (1 - \alpha) r z \mathbb{P}(\text{Bi}(r-1, 1-z) \leq k-1), \end{aligned}$$

and

$$h_1(z; \alpha, \mathbf{p}(r), 1) = \mathbb{P}(\text{Bi}(r, 1-z) \leq k-1).$$

We see that if $\alpha > \alpha_c$, with

$$\alpha_c = 1 - \inf_{0 < z \leq 1} \frac{z}{\mathbb{P}(\text{Bi}(r-1, 1-z) \leq k-1)},$$

then $\hat{z}(\mathbf{p}(r), \alpha, 1) = 0$ so that by Theorem 4, the epidemic reaches all nodes (asymptotically). This result was first proved in [1], see also Theorem 5.1 in [14] which corresponds exactly to our Theorem 4 in this particular setting.

5. PROOF OF THE MAIN RESULTS

We first give an overview of the proof of our main results and then give the technical details of the proofs.

5.1 Sketch of the proof

It is well-known that it is often simpler to study the random multigraph $G^*(n, \mathbf{d})$ with given vertex sequence $\mathbf{d} = (d_i)_1^n$ defined in Section 2.1. We consider asymptotics as the numbers of vertices tend to infinity and thus assume throughout the paper that we are given, for each n , a sequence $\mathbf{d}^{(n)} = (d_i^{(n)})_1^n$ with $\sum_i d_i^{(n)}$ even. For notational simplicity we will usually not show the dependency on n explicitly. We may obtain $G(n, \mathbf{d})$ by conditioning the multigraph $G^*(n, \mathbf{d})$ on being a simple graph. By [13], we know that the condition $\sum_i d_i^2 = O(n)$ implies $\liminf \mathbb{P}(G^*(n, \mathbf{d}) \text{ is simple}) > 0$. In this case, many results transfer immediately from $G^*(n, \mathbf{d})$ to $G(n, \mathbf{d})$, for example, every result of the type $\mathbb{P}(\mathcal{E}_n) \rightarrow 0$ for some events \mathcal{E}_n , and thus every result saying that some parameter converges in probability to some non-random value. This includes every result in the present paper.

We will in this paper study the random multigraph $G^*(n, \mathbf{d})$ and in a last step (left to the reader) transfer the results to $G(n, \mathbf{d})$ by conditioning. More precisely, we define the graph $G^*(n, \mathbf{d}, \mathbf{k})$, i.e. we add a label to each vertex corresponding to its threshold and say that it is the configuration model associated to the degree-threshold sequence (\mathbf{d}, \mathbf{k}) and asymptotic degree-threshold distribution $\mathbf{p} = (p_s t_{s\ell})_{s\ell}$.

The proof is now done in two steps:

- first, as explained in Section 2.3, we need to remove from the graph $G^*(n, \mathbf{d}, \mathbf{k})$ the vaccinated nodes. We denote the resulting graph $G^*(n, \mathbf{d}, \mathbf{k})_{\beta^p}$ for the degree based vaccination and $G^*(n, \mathbf{d}, \mathbf{k})_{\beta^A}$ for the acquaintance vaccination. Our proof is inspired by ideas from

[10, 14]: in order to derive results for these percolation models we show that they are again configuration models but for different degree-threshold sequence. This is done in Sections 5.3 and 5.4, in particular we compute the modified asymptotic degree-threshold distribution for the two graphs $G^*(n, \mathbf{d}, \mathbf{k})_{\beta^p}$ and $G^*(n, \mathbf{d}, \mathbf{k})_{\beta^A}$.

- then, we run the algorithm of Section 2.2 on a general graph $G^*(n, \mathbf{d}, \mathbf{k})$ in order to compute the final size of the epidemic in a similar way as in [15]. The results for $G^*(n, \mathbf{d}, \mathbf{k})_{\beta^p}$ and $G^*(n, \mathbf{d}, \mathbf{k})_{\beta^A}$ follow directly from this general result applied to the modified degree-threshold distribution. Due to the lack of space, we omitted this last step and refer to [17] for a proof with full details.

5.2 A Lemma for death processes

A pure death process with rate 1 is a process that starts with some number of balls whose lifetime are i.i.d. rate 1 exponentials. We modify the life-time of the balls as follows: for each ball when it normally dies, with probability π we set it as its real dead and with probability $1 - \pi$ we set the dead to ∞ . We call such a death process a death process with immortal balls. Now consider n bins with independent death processes with immortal balls. To each bin, we attach a couple (d_i, k_i) where d_i is the number of balls at time 0 in the bin and $k_i = K_i(d_i)$ is the corresponding threshold. Let $N_j^{(n)}(t)$ denote the number of balls in bin j at time t , where $j = 1, \dots, n$ and $t \geq 0$. Let $n_{s\ell} = |\{i : d_i = s, K_{d_i} = \ell\}|$ and let $U_{sr,\ell}^{(n)}(t)$ be the number of bins that have s balls at time 0 and r balls at time t and threshold ℓ , $U_{sr,\ell}^{(n)}(t) = |\{i, N_i^{(n)}(t) = r, d_i = s, K_{d_i} = \ell\}|$. In particular $\sum_r U_{sr,\ell}(0) = U_{ss,\ell}(0) = n_{s\ell}$. In what follows we suppress the superscripts to lighten the notation. The following Lemma is an extension of Lemma 4.4 of [15]:

LEMMA 10. *Consider n independent death process with immortal balls $N_i(t)$ such that $n_{s\ell}/n \rightarrow p_{s\ell}$, where $p_s := \sum_{\ell} p_{s\ell}$ satisfies Condition 1. Then, with the above notation, as $n \rightarrow \infty$,*

$$\sup_{t \geq 0} \sum_{s,\ell} \sum_{r=0}^s r |U_{sr,\ell}(t)/n - p_{s\ell} b_{sr}(1 - \pi + \pi e^{-t})| \rightarrow^p 0. \quad (10)$$

PROOF. First fix integers $j \leq s$. Consider the $n_{s\ell}$ bins that start with s balls (and with threshold ℓ). For $i = 1, \dots, n_{s\ell}$, let T_i be the time the j th ball is removed from the i th such bin. Then $|\{i : T_i \leq t\}| = \sum_{r=0}^{s-j} U_{sr,\ell}(t)$. Moreover, the number of balls remaining in one of these bins at time t has a Binomial distribution with parameters s and $\mathbb{E}[e^{-tB}] = 1 - \pi + \pi e^{-t}$ (where B is a Bernoulli r.v. with parameter π). Thus $\mathbb{P}(T_i \leq t) = \sum_{r=0}^{s-j} b_{sr}(1 - \pi + \pi e^{-t})$. Hence we have

$$\sup_{t \geq 0} \left| \frac{1}{n} \sum_{r=0}^{s-j} U_{sr,\ell}(t) - \frac{n_{s\ell}}{n} \sum_{r=0}^{s-j} b_{sr}(1 - \pi + \pi e^{-t}) \right| \rightarrow^p 0.$$

Since $n_{s\ell}/n \rightarrow p_{s\ell}$ we have for all $j, s \geq 0$:

$$\sup_{t \geq 0} \left| \frac{1}{n} \sum_{r=0}^{s-j} U_{sr,\ell}(t) - p_{s\ell} \sum_{r=0}^{s-j} b_{sr}(1 - \pi + \pi e^{-t}) \right| \rightarrow^p 0.$$

In particular each term in (10) tends to 0 in probability.

Let $\epsilon > 0$ and S be such that $\sum_{s=S}^{\infty} sp_s < \epsilon$. By Condition (iii) of Condition 1, we have $\sum_{s,\ell} sn_{s\ell}/n \rightarrow \lambda = \sum_s sp_s$.

Hence, also $\sum_{s \geq S} \sum_{\ell} sn_{s\ell}/n \rightarrow \sum_{s=S}^{\infty} sp_s < \epsilon$. So that for sufficiently large n , we get $\sum_{s \geq S} \sum_{\ell} sn_{s\ell}/n < \epsilon$ and

$$\begin{aligned} & \sup_{t \geq 0} \sum_{s \geq S} \sum_{\ell} \sum_{r=0}^s r |U_{sr,\ell}(t)/n - p_{s\ell} b_{sr}(1 - \pi + \pi e^{-t})| \\ & \leq \sup_{t \geq 0} \sum_{s \geq S} \sum_{\ell} \sum_{r=0}^s r (U_{sr,\ell}(t)/n + p_{s\ell} b_{sr}(1 - \pi + \pi e^{-t})) \\ & \leq \sum_{s \geq S} \sum_{\ell} s(k_{s\ell}/n + p_{s\ell}) < 2\epsilon. \end{aligned}$$

Hence (10) holds. \square

5.3 Degree based vaccination

The degree based vaccination corresponds exactly to the site percolation described in [14], but as explained in Section 5.1, we need to keep track of the mark \mathbf{k} . More precisely, we do the deletions in two steps as in [14]: for each vertex i , replace it with probability $\beta_{d_i}^P$ by d_i new vertices of degree 1; we further colour the new vertices red. Then clean up by removing all red vertices. Note that the (random) explosions change the number of vertices, but not the number of half-edges. Moreover, given the set of explosions, there is a one-to-one correspondence between configurations before and after the explosions, and thus, if we condition on the new degree, threshold sequence, the exploded graph is still described by the configuration model. Furthermore, by symmetry, when removing the red vertices, all vertices of degree 1 are equivalent (note that the threshold of a degree 1 node is always 0), so we may just as well remove the right number of vertices of degree 1, but choose them uniformly at random. Hence we can obtain $G^*(n, \mathbf{d}, \mathbf{k})_{\beta^P}$ as follows: for each vertex i , replace it with probability $\beta_{d_i}^P$ by d_i new vertices of degree 1. Let $(\mathbf{d}_{\beta^P}, \mathbf{k}_{\beta^P})$ be the resulting (random) degree-threshold sequence. Note that \mathbf{k}_{β^P} is obtained from \mathbf{k} by adding a zero threshold to all new degree 1 vertices. Let \tilde{n} be the length of \mathbf{d}_{β^P} and let n_+ be the number of new vertices. Construct the random graph $G^*(\tilde{n}, \mathbf{d}_{\beta^P}, \mathbf{k}_{\beta^P})$. Finish by deleting n_+ randomly chosen vertices of degree 1.

The degree sequence \mathbf{d}_{β^P} is computed in Section 2.1 of [14]. Let $\tilde{n}_{j\ell} = |\{i : \tilde{d}_i = j, K_i(d_i) = \ell\}|$ the number of vertices in $G^*(\tilde{n}, \mathbf{d}_{\beta^P}, \mathbf{k}_{\beta^P})$ with degree j and threshold ℓ . We have $\tilde{n} \sim \zeta n$ where $\zeta = \sum_{j=0}^{\infty} (1 - \beta_j^P + j\beta_j^P) p_j$ and

$$n_+ \sim \sum_{j=1}^{\infty} j\beta_j^P p_j n \quad (11)$$

$$\tilde{n}_{j\ell} \sim (1 - \beta_j^P) p_j t_{j\ell} n, \quad j \neq 1 \quad (12)$$

$$\tilde{n}_{10} \sim \left((1 - \beta_1^P) p_1 + \sum_{j=1}^{\infty} j\beta_j^P p_j \right) n, \text{ and,} \quad (13)$$

$$\tilde{n}_{1\ell} = 0 \text{ for } \ell \neq 0. \quad (14)$$

We now consider the deletion process as follows: start with all half-edges white. Color red n_+ half-edges of degree 1. Each half-edge has a lifetime which is exponentially distributed with mean 1. Each time an half-edge dies spontaneously, we kill a living red half-edge. We stop when there is no more living red half-edge. Let $A(t)$ be the number of living white half-edges at time t and $B(t)$ be the number of living red half-edges.

We have

$$A(t) + B(t) \sim n\lambda e^{-2t},$$

and

$$A(t) = \sum_{s \geq 1, r, \ell \geq 0} r U_{sr,\ell}(t),$$

where the processes $U_{s\ell,r}$ are as in Lemma 10 with $p_{s\ell} = \lim \frac{\tilde{n}_{s\ell}}{\tilde{n}}$ for $s \neq 1$ and $p_{10} = (1 - \beta_1^P) p_1 / \zeta$. Hence we have

$$\begin{aligned} \frac{B(t)}{n} & \sim \lambda e^{-2t} - (1 - \beta_1^P) p_1 e^{-t} - \sum_{s \geq 2, r \geq 0} r(1 - \beta_s^P) p_s b_{sr}(e^{-t}) \\ & = \lambda e^{-2t} - \left(\sum_{s \geq 1} s(1 - \beta_s^P) p_s \right) e^{-t}, \end{aligned}$$

so that $B(\tau) \sim 0$ iff $e^{-\tau} = \frac{1}{\lambda} \sum_{s \geq 1} s(1 - \beta_s^P) p_s = x$.

Note that the number of nodes in the final graph is

$$n^P = \tilde{n} - n_+ \sim \left(\sum (1 - \beta_j^P) p_j \right) n.$$

Hence we proved

LEMMA 11. *The random graph obtained after deletion of the vaccinated nodes in the degree based vaccination with parameter β^P is a configuration model with $n^P = \zeta^P n + o_P(n)$ nodes and*

$$p_{j\ell}^P = \frac{1}{\zeta^P} \sum_{s \geq j} (1 - \beta_s^P) p_s t_{s\ell} b_{sj}(x),$$

where $\zeta^P = \sum (1 - \beta_j^P) p_j$, $x = \frac{1}{\lambda} \sum_{s \geq 1} s(1 - \beta_s^P) p_s$.

5.4 Acquaintance vaccination

We consider the acquaintance vaccination and prove the following lemma

LEMMA 12. *The random graph obtained after deletion of the vaccinated nodes in the acquaintance vaccination with parameter β^A is a configuration model with $n^A = \zeta^A n + o_P(n)$ nodes and*

$$p_{j\ell}^A \rightarrow \frac{1}{\zeta^A} \sum_{s \geq j} y^s p_s t_{s\ell} b_{sj}((1 - \beta_s^A) \hat{p}),$$

where $y = \frac{1}{\lambda} \sum_{s \geq 1} s(1 - \beta_s^A) p_s$, $\zeta^A = \sum_{\ell} p_{\ell} y^{\ell}$ and $\hat{p} = \frac{\sum_{\ell} \ell p_{\ell} y^{\ell} (1 - \beta_{\ell}^A)}{\lambda y^2}$.

PROOF. The proof proceeds in two phases: we first study the 'propagation' of the vaccination and then remove the vaccinated nodes.

Phase 1: we start with all half-edges white. We color red each half-edge of degree k with probability β_k^A . A red half-edge will propagate the inoculation to the vertex it will be associated with. We now study this propagation process using standard coupling methods. Each half-edge has a lifetime which is exponentially distributed with mean 1. Each time an half-edge dies spontaneously, we kill a living red half-edge (meaning that these two half-edges are paired). We stop when there is no more living red half-edge. Let $A(t)$ be the number of living white half-edges at time t and $B(t)$ be the number of living red half-edges.

As in previous section (or in [14]), we can compute $n_{\ell j}$ the number of vertices of degree ℓ with j white edges:

$$n_{\ell j} \sim b_{\ell j}(1 - \beta_{\ell}^A) p_{\ell} n.$$

By an easy application of Lemma 10, we have

$$\begin{aligned}
A(t) &\sim \sum_{\ell \geq j \geq r} r n_{\ell j} b_{jr}(e^{-t}) \\
&\sim n \sum_{\ell \geq r} r p_{\ell} \sum_{j \geq r} b_{\ell j} (1 - \beta_{\ell}^A) b_{jr}(e^{-t}) \\
&= n \sum_{\ell \geq r} r p_{\ell} b_{\ell r} ((1 - \beta_{\ell}^A) e^{-t}) \\
&= n \sum_{\ell} \ell (1 - \beta_{\ell}^A) e^{-t} p_{\ell} = n \lambda e^{-t} y,
\end{aligned}$$

with $y = \sum_{\ell} \ell (1 - \beta_{\ell}^A) \frac{p_{\ell}}{\lambda}$.

By considering the death process without taking into account the colors, we get

$$A(t) + B(t) \sim n \lambda e^{-2t}.$$

Hence we have $\frac{B(t)}{n} \sim \lambda e^{-2t} - e^{-t} \lambda y$ and the process stop when $B(\tau) = 0$, so that we have $\tau \sim -\log y$.

At the end of this phase 1, we have $A(0) - A(\log y) \sim n \lambda y (1 - y)$ dead of white half-edges. All these deaths are spontaneous hence the number of spontaneous dead of red half-edges D is such that $n \lambda y (1 - y) + 2D \sim B(0)$, so that, $D \sim \frac{n \lambda (1 - y)^2}{2}$.

We now add a color: each time a red half-edge dies spontaneously, it is paired with another red half-edge. We color both these half-edges blue. Clearly the number of blue half-edges is twice the number of spontaneous dead of red half-edges and hence is $n \lambda (1 - y)^2 + o_p(n)$. Since the number of red half-edges at the beginning of phase 1 is $n \lambda (1 - y) + o_p(n)$, we have: given the fact that an half-edge is red at the start of phase 1, it has probability $(1 - y) + o(1)$ to be blue at the end of the phase 1 with this new coloring scheme. Let $\tilde{n}_{\ell j}$ be the number of degree ℓ vertices with j white edges at the beginning of phase 1 and no blue half-edges at the end of phase 1. We have $\tilde{n}_{\ell j} \sim n p_{\ell} b_{\ell j} (1 - \beta_{\ell}^A) y^{\ell - j}$. Note also that the total number of deaths is equal to the number of deaths of white half-edges $n \lambda y (1 - y) + o(n)$ plus the number of red half-edges $n \lambda (1 - y) + o(n)$, hence the number of living edges at the end of phase 1 is $n \lambda y^2 + o(n)$.

Phase 2: we start from the end of phase 1. A vertex with white half-edges which are all alive (i.e. not paired yet) and no blue half-edges is colored green. All other vertices (and corresponding half-edges) are colored black. The interpretation is as follows: all black vertices have been paired with at least one red half-edge and hence correspond to the vaccinated nodes. The green vertices have not been paired to a red half-edge and are not vaccinated. Note however that all red half-edges are by definition dead (i.e. have been paired in phase 1). Hence the red half-edges of the green vertices are paired to vaccinated nodes. We now remove all vaccinated nodes and their incident edges in continuous time in a similar fashion as in phase 1: each time a live half-edge dies, we kill a living black half-edge.

Let $C(t)$ be the number of living green half-edges and $D(t)$ the number of living black edges. Since the number of living half-edges at the end of phase 1 is $n \lambda y^2 + o(n)$, we have

$$C(t) + D(t) \sim n \lambda y^2 e^{-2t}.$$

By a standard application of Lemma 10, we have

$$C(t) \sim \sum_{\ell \geq j \geq r} r \tilde{n}_{\ell j} y^j b_{jr}(e^{-t}),$$

because the number of degree ℓ green vertex with j living edges at the end of phase 1, is $\tilde{n}_{\ell j} b_{jr}(y) + o(n) = \tilde{n}_{\ell j} y^j + o(n)$. Hence we get

$$\begin{aligned}
C(t) &\sim n \sum_{\ell \geq j} p_{\ell} b_{\ell j} (1 - \beta_{\ell}^A) y^{\ell} \sum_{r \leq j} r b_{jr}(e^{-t}) \\
&= n e^{-t} \sum_{\ell} p_{\ell} y^{\ell} \sum_{j \leq \ell} j b_{\ell j} (1 - \beta_{\ell}^A) \\
&= n e^{-t} \sum_{\ell} \ell p_{\ell} y^{\ell} (1 - \beta_{\ell}^A).
\end{aligned}$$

Let

$$\hat{p} = \frac{\sum_{\ell} \ell p_{\ell} y^{\ell} (1 - \beta_{\ell}^A)}{\lambda y^2},$$

so that the process stop when $D(\tau) = 0$ with $\tau \sim -\log \hat{p}$.

As explained above, the number of vaccinated vertices is the number of black vertices:

$$v \sim n - n \sum_{\ell \geq j} p_{\ell} b_{\ell j} (1 - \beta_{\ell}^A) y^{\ell} = n (1 - \zeta^A),$$

with the notation of the lemma.

At the end of phase 2, the dead half-edges (which are paired by definition) correspond exactly to the edges adjacent to a vaccinated graph and should be removed from the graph. Clearly, once these edges are removed, the graph is still a configuration model and we compute now the degree-threshold sequence.

Let $\tilde{n}_{\ell r}$ be the number of vertices with degree ℓ in the initial graph and degree r in the final graph (i.e. having r living half-edges at the end of phase 2), then we have

$$\begin{aligned}
\tilde{n}_{\ell r} &\sim \sum_{j \geq r} \tilde{n}_{\ell j} y^j b_{jr}(\hat{p}) \\
&\sim n p_{\ell} y^{\ell} b_{\ell r} ((1 - \beta_{\ell}^A) \hat{p}).
\end{aligned}$$

The number of remaining vertices is $\tilde{n} = n - v \sim n \sum_{\ell} p_{\ell} y^{\ell}$. The statement of the lemma follows from: $p_{j\ell}^A \sim \frac{\sum_{s \geq j} \tilde{n}_{s j} t_{s\ell}}{\tilde{n}}$.

□

6. SUMMARY

We have introduced a new model for the spread of epidemics combining neighborhood effects and random transmission: the percolated threshold model. For a degree based attack, we gave a rigorous analysis of the final size of the epidemic on a random graph with specified degree distribution. We also analyzed rigorously the impact of a degree based vaccination or a acquaintance vaccination. Our main theorems give new results for the condition under which a single node can trigger a large cascade (the cascade condition). They allow to quantify the impact of the different vaccinations and allow to define a security metric which allow to compare them. In particular, we show that the acquaintance vaccination proposed in [7] has the advantage of requiring no global information and allows a much better protection of the network than the uniform vaccination in all cases.

7. REFERENCES

- [1] J. Balogh and B. G. Pittel. Bootstrap percolation on the random regular graph. *Random Structures Algorithms*, 30(1-2):257–286, 2007.

- [2] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. On the spread of viruses on the internet. In *SODA*, pages 301–310. SIAM, 2005.
- [3] B. Bollobás. *Random graphs*. Cambridge University Press, Cambridge, second edition, 2001.
- [4] C. Borgs, J. Chayes, A. Ganesh, and A. Saberi. How to distribute antidote to control epidemics.
- [5] T. Britton, S. Janson, and A. Martin-Löf. Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Adv. in Appl. Probab.*, 39(4):922–948, 2007.
- [6] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin. Resilience of the internet to random breakdowns. *Phys. Rev. Lett.*, 85(21):4626–4628, Nov 2000.
- [7] R. Cohen, S. Havlin, and D. ben Avraham. Efficient immunization strategies for computer networks and populations. *Phys. Rev. Lett.*, 91(24), 2003.
- [8] M. Draief, A. Ganesh, and L. Massoulié. Thresholds for virus spread on networks. *Ann. Appl. Probab.*, 18(2):359–378, 2008.
- [9] R. Durrett. *Random graph dynamics*. Cambridge University Press, Cambridge, 2007.
- [10] N. Fountoulakis. Percolation on sparse random graphs with given degree sequence, 2007.
- [11] A. J. Ganesh, L. Massoulié, and D. F. Towsley. The effect of network topology on the spread of epidemics. In *INFOCOM*, pages 1455–1466, 2005.
- [12] M. O. Jackson and L. Yariv. Diffusion of behavior and equilibrium properties in network games. *The American Economic Review*, 97(2), 2007.
- [13] S. Janson. The probability that a random multigraph is simple, arxiv.org:math/0609802, 2006.
- [14] S. Janson. On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*, 14:86–118, 2009.
- [15] S. Janson and M. J. Luczak. A simple solution to the k -core problem. *Random Structures Algorithms*, 30(1-2):50–62, 2007.
- [16] J. Kleinberg. Cascading behavior in networks: algorithmic and economic issues. In *Algorithmic game theory*, pages 613–632. Cambridge Univ. Press, Cambridge, 2007.
- [17] M. Lelarge. Diffusion and cascading behavior in random networks, <http://www.di.ens.fr/~lelarge/>.
- [18] M. Lelarge. Diffusion of innovations on random networks: Understanding the chasm. In *WINE*, pages 178–185, 2008.
- [19] M. Lelarge. Economics of Malware: Epidemic Risks Model, Network Externalities and Incentives. In *Fifth bi-annual Conference on The Economics of the Software and Internet Industries*, 2009.
- [20] M. Lelarge and J. Bolot. A local mean field analysis of security investments in networks. In *Proceedings of ACM SIGCOMM NetEcon '08*, pages 25–30, 2008.
- [21] M. Lelarge and J. Bolot. Network externalities and the deployment of security features and protocols in the internet. In *Proceedings of ACM SIGMETRICS*, pages 37–48, 2008.
- [22] P. Maniatis, M. Roussopoulos, T. J. Giuli, D. S. H. Rosenthal, and M. Baker. The lockss peer-to-peer digital preservation system. *ACM Trans. Comput. Syst.*, 23(1):2–50, 2005.
- [23] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures Algorithms*, 6(2-3):161–179, 1995.
- [24] S. Morris. Contagion. *Rev. Econom. Stud.*, 67(1):57–78, 2000.
- [25] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev.*, 45(2):167–256 (electronic), 2003.
- [26] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical review letters*, 86(14):3200–3203, 2001.
- [27] D. Peleg. Local majorities, coalitions and monopolies in graphs: a review. *Theoret. Comput. Sci.*, 282(2):231–257, 2002.
- [28] O. Riordan. The k -core and branching processes. *Combin. Probab. Comput.*, 17(1):111–136, 2008.
- [29] F. Vega-Redondo. *Complex social networks*, volume 44 of *Econometric Society Monographs*. Cambridge University Press, Cambridge, 2007.
- [30] D. J. Watts. A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci. USA*, 99(9):5766–5771 (electronic), 2002.
- [31] N. Weaver, V. Paxson, S. Staniford, and R. Cunningham. A taxonomy of computer worms. In *WORM '03: Proceedings of the 2003 ACM workshop on Rapid malware*, pages 11–18, New York, NY, USA, 2003. ACM.
- [32] C. C. Zou, L. Gao, W. Gong, and D. Towsley. Monitoring and early warning for internet worms. In *CCS '03: Proceedings of the 10th ACM conference on Computer and communications security*, pages 190–199, New York, NY, USA, 2003. ACM.